The Report Committee for Wonsun Ryu
Certifies that this is the approved version of the following Report:

The Relationship between Institutional Characteristics and Six-Year Graduation Rates: Multilevel Modeling for Change

APPROVED BY
SUPERVISING COMMITTEE:

Daniel A. Powers, Supervisor

Thomas W. Sager

The Relationship between Institutional Characteristics and Six-Year Graduation Rates: Multilevel Modeling for Change

by

Wonsun Ryu

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Statistics

The University of Texas at Austin

May 2019

Dedicated to Almighty God

Acknowledgements

I cannot thank my supervisor Dr. Daniel Powers enough for his dedicated supervision and unwavering support throughout this study. I would also like to express my sincere appreciation to Dr. Thomas Sager, reader of this report, for his valuable feedback and extended support throughout the writing process of this report. Last but not the least, I wish to express my deepest gratitude to my family – my wife, my two daughters, and my parents – for their unconditional love and endless support all the time. This report would have never been possible without their dedication and support.

Abstract

The Relationship between Institutional Characteristics and Six-Year Graduation Rates: Multilevel Modeling for Change

Wonsun Ryu, M.S. Stat.

The University of Texas at Austin, 2019

Supervisor: Daniel A. Powers

A number of studies have explored the relationship between institutional characteristics (e.g., institutional type, enrollment size, institutional expenditures) and graduation rates as a key indicator to measure institutional performance. However, researchers paid limited attention to investigate the relationship between institutional characteristics and growth trajectories of graduation rates. Therefore, this study explores longitudinal patterns of graduation rates and the predictability of these patterns on the basis of institutional characteristics by using the Integrated Postsecondary Education Data System (IPEDS) data and applying multilevel modeling for change. By fitting a growth model to the IPEDS data on 2011-2017 graduation rates, the results indicate that public four-year institutions' graduation rates increased more rapidly than private, not-for-profit four-year colleges and universities albeit lower graduation rates in 2011. In addition, the findings indicate that annual changes in the graduation rate did not differ between doctoral or research-oriented colleges and universities and other types of four-year institutions while doctoral or research

institutions had higher graduation rates in 2011. In addition, when time-varying predictors were added to the growth model, significant relationships among institutional expenditures, enrollment size, the percentage of federal grant recipients, and graduation rates were found.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction and Brief Literature Review

Graduation rates are widely used to measure institutional performance in higher education (Burke & Minassians, 2002; Christal, 1998; Shin, 2010). Institutional performance can be also measured in other ways, such as student retention, student satisfaction, job placement rate, research productivity, and campus diversity; however, many of these measures are ultimately designed to increase the number of students who complete a degree (Tandberg & Hillman, 2014). Federal and state governments have emphasized the need for responsibility of colleges and universities to increase or maintain graduation rates since the 1990s (National Research Council, 2013) because higher education provides considerable value to individuals, the economies where educated individuals work and live, and society in general. Institutional leaders also sought to understand if they were operating institutions efficiently and effectively to enhance student outcomes.

By using graduation rates as a performance measure, researchers have explored how institutional characteristics (e.g., institutional type, enrollment size) are associated with graduation rates (e.g., Ishitani, 2006; Oseguera, 2005; Titus, 2006a, 2006b; Zhan et al., 2018). Although prior studies have provided empirical insights into building predictive models to improve institutional graduation rates, existing research paid little attention to understanding the association between institutional characteristics and changes in the gradual graduation rates. In other words, much less is known about the temporal nature of graduation rate. Therefore, it is critical to explore how institutional characteristics are associated with patterns of change in graduation rates over time. In particular, it is critical to capture the dynamic processes of institutional expenditures that are subject to gradual changes.

This study explores longitudinal patterns of graduation rates and the predictability of these patterns on the basis of institutional characteristics. It seeks to answer the following research question: what institutional characteristics are associated with patterns of change in graduation rates? In addition, this study intends to address statistical methodologies which are appropriate to investigate longitudinal graduation patterns. Specifically, this study utilizes multilevel modeling for change, which allows me to investigate change of graduation rates over time. I build a statistical model based on prior research, model testing, and comparison of model fit in order to address my research question.

BRIEF LITERATURE REVIEW

Using institutional-level data such as IPEDS, researchers have explored how institutional characteristics are associated with degree production. Researchers found that students who attend public institutions are less likely to graduate than students who are in private colleges and universities (Ishitani, 2006; Oseguera, 2005; Titus, 2006a, 2006b; Zhan Hu, & Sensenig, 2018). Researchers used the Carnegie classification to identify comparable groups of institutions, such as Doctoral/Research, Master's, Bachelor's, etc., researchers sought to explore how an institution's mission is associated with its graduation rate. Cragg (2009) estimated the probabilities of graduation by Carnegie classification and found that institutional classification was not associated with degree completion. In other words, there were no differences between bachelor's and other institutional categories such Doctoral/Research, Master's, etc. However, Zhang, Hu, and Sensenig (2013) found that research and doctorate-granting public universities had an impact on baccalaureate degree completion while other types of institutions (e.g., Master's public, Bachelor's public) had no effects. Gross, Torres, & Zerquera (2013) and Gross & Berry (2016) also reported that students in research universities have higher probability of degree completion than those in state, urban, or regional institutions. Regarding institutional enrollment size, Oseguera (2005) and Titus (2006a) reported that enrollment size is negatively associated with degree completion while Cragg (2009) found no relationship between enrollment size and degree completion. Sanford and Hunter (2011) reported the percentage of Pell recipient students of color is related to degree completion while the percentage of Pell recipients are negatively associated with completion.

Institutional expenditures are also important sources of influences on student graduation. Titus (2006a, 2006b) reported that only educational and general expenditures per student increased the probability of a student's graduation after controlling for all other factors. Consistent with the findings of prior studies (Titus, 2006a, 2006a), Goenner and Snaith (2004) found that two-year graduation rate is positively associated with educational and general expenditures while four- and five-year graduation rates are not related to those expenditures. Gansemer-Topf & Schuh (2003) found that expenditures on instruction and academic support increased institutional graduation rates. Gansemer-Topf & Schuh (2006) extended their study by including additional variables on expenditures and reported that many types of expenditures (i.e., instruction, academic support, institutional support, and grants) are important to increase graduation rates while student services expenditures were not associated with degree production. However, the more recent study by Webber and Ehrenberg (2010) found student services are important in degree production whereas academic support has no significant effect on institutional graduation rates.

In summary, previous studies have shown that institutional characteristics can play a role in predicting degree completion, but results are mixed outcomes. This collection of studies offers valuable strategies for inclusion institutional characteristics as control measures in the analytic model. Therefore, this study includes institutional control (i.e., public versus private, not-for-profit), Carnegie classification, enrollment size, percentage of federal grant recipients, and five forms of expenditures: 1) instructions, 2) research, 3) academic support, 4) student services, and 5) institutional support – which are found in the previous studies reviewed.

Chapter Two: Method

DATA SOURCES

The purpose of this study is to test a model of institutional performance in graduation rates for four-year, public and private, not-for-profit, Title IV institutions in the United States. Data for this study come from the Integrated Postsecondary Education Database System administered by National Center for Education Statistics (IPEDS). IPEDS data is longitudinal and institutional-level data on accredited postsecondary institutions that are eligible for Title IV financial aid. To model change, longitudinal data that describe how each person in the sample changes over time. IPEDS collects data on postsecondary education in the United States in the following areas: institutional characteristics, institutional prices, admissions, enrollment, student financial aid, degrees and certificates conferred, student persistence and success (retention rates, graduation rates, and outcome measures), institutional human resources, fiscal resources, and academic libraries. IPEDS data are available to the public through the IPEDS Website, http://nces.ed.gov/ipeds/.

Four-year degree-granting institutions in the United States constitute the population of interest for this study. I use six-year graduation rate data from IPEDS for each year between 2011 and 2017 which covers the 2005 to 2011 cohorts. Colleges and universities report to IPEDS their graduation rates for fall semester cohorts of first-time, full-time students, which represent less than half of the college students in the United States (Schneider & Yin, 2011). In addition, only institutions with valid information on all variables were determined to be eligible for the analytic sample. Consequently, the data collection and cleaning certainly miss a large share of the college-going student population and degree-granting institutional population. The final sample of 1,260 four-year public and private, not-for-profit degree-granting institutions in the United States is included in this present study.

VARIABLES

In this study, the dependent variable of interest is graduation rate. The time-varying dependent variable allows me to assess both initial systemic differences between institutions and systematical change over time. The time variable year is added to the analysis and recorded as year number from the starting point from 2011 through 2017, coded 0 through 6, respectively. Selecting time-invariant independent variables driven by the literature review, I first include institution type (e.g., public versus private, not-for-profit) and Carnegie classification to account for inter-institution variance in changes in institutional graduation rate. Values for these time-invariant variables are the same no matter when they are observed. Informed by the preceding literature review of past studies on six-year graduation rates, I also include time-varying variables such as enrollment size, percentage of students awarded federal grant aid, and five forms of institutional expenditures. Table 1 presents the definitions of all the variables used in the analysis.

ANALYSIS

Due to the focus on the patterns of longitudinal change in six-year graduation rates at the institutional-level, this study utilizes multilevel modeling for change to explore the relationship between institutional characteristics such as expenditures and six-year graduation rates and examines the effects of institutional characteristics on change of six-year graduation rates over time. While traditional statistical analytic approaches can be biased when multiple observations are collected from subjects over time, multilevel modeling for change is designed to address dependency in data such as autocorrelation and heteroscedasticity (West, Ryu, Kwok, & Cham, 2011). In multilevel models for change, data are structured in two hierarchical levels –within-institution (Level 1) and between-institution (Level 2). Thus, multilevel modeling for change allows researchers to explore individual changes or growth trajectories on the variable of interest and differences between individuals (Singer & Willett, 2003).

Table 1: Definition of Variables

Variable	Definition and Coding
Dependent Variable	
Graduation Rate	6-year baccalaureate degree completion rate from 2011 to 2017
Independent Variable	
Time-Invariant	
Public	1 = Public, 0 = Private, not-for-profit
Doctoral/Research	1 = Doctoral/Research, 0 = Not Doctoral/Research
Time-Varying	
FTE enrollment	Full-time equivalent fall enrollment (in 1000s) from 2005 to 2011 (*lagged)
% federal grant aid	Percentage of full-time first-time undergraduates awarded federal grant aid from 2005 to 2011 (*lagged)
Instruction expenditures	Instruction expenditures per FTE (in $1000s) from 2005 to 2011 (*lagged)
Research expenditures	Research expenditures per FTE (in $1000s) from 2005 to 2011 (*lagged)
Academic support expenditures	Academic support expenditures per FTE (in $1000s) from 2005 to 2011 (*lagged)
Student services expenditures	Student services expenditures per FTE (in $1000s) from 2005 to 2011 (*lagged)
Institutional support expenditures	Institutional support expenditures per FTE (in $1000s) from 2005 to 2011 (*lagged)

This study applies a four-step process using Singer and Willett’s (2003) exploratory modeling strategy: (1) unconditional means model, (2) unconditional growth model, (3) conditional growth model with time-invariant predictors, and (4) conditional growth model with time-invariant and time-varying predictors. The unconditional means model is the preliminary check on whether multilevel modeling for change is appropriate for this analysis by partitioning the total variation in the outcome variable (i.e., six-year graduation rates). Second, the unconditional growth model intends to assess the effects of time (i.e., year) on change of graduation rates and also determine whether there is significant variance to be explained in graduation rate change from institutional characteristics. In addition, two separate conditional growth models are introduced to examine whether institutional characteristics are significant predictors of change in graduation rates. Two time-invariant predictors (i.e., Public, Doctoral/Research) are added into the unconditional growth model and then the remaining time-varying predictors of institutional characteristics are included to build the full model. Lastly, I explore different error covariance structures to address heteroscedasticity and autocorrelation in my longitudinal data. All results were obtained using the STATA statistical analysis software, version 15.1.

Chapter Three: Results

DESCRIPTIVE STATISTICS

Table 2 presents descriptive statistics of all the variables used in the analysis. From 2011 to 2017, the graduation rates on average have slightly increased from 55.66% to 56.78% with standard deviations of 17.88 and 17.67, respectively. In the sample, 34% of all institutions are public 4-year institutions. Only 17% of four-year institutions are classified the doctoral/research institutional group. The percentage of students who received federal grant aid has sharply increased by 9.36% from 2011 to 2017 while the increase in full-time enrollment is relatively small. The population average institutional expenditures have generally increased over time. Research expenditures particularly have considerably larger standard deviations than other types of expenditures, indicating research expenditures vary across institutions in the sample.

Before testing multilevel models for change, I began fitting a separate linear change model to each institution's empirical graduation growth rate in order to illustrate how the initial status and patterns of change within six-year graduation rates differ across institutions. Figure 1 presents the distribution of fitted intercepts and slopes. The results show that the distribution of fitted slopes is denser than fitted intercepts while both distributions of fitted intercepts and slopes appeared to be approximately normal. The fitted intercepts and slopes are summarized in Table 3. Further, average estimated values of intercepts and slopes are 55.43 and 0.21, respectively. The results suggest that the average institution in the sample has an observed 55.43% of six-year graduation rate in 2011 and the six-year graduation rate increased by an estimate of 0.21% per year. As presented Figure 1 and Table 3, standard deviations, minimum and maximum show that institutions are scattered widely around these averages. Accordingly, the results indicate that institutions differ considerably in their fitted initial status and fitted rates of change. Finally, the negative correlation between initial status and rate of change suggests that institutions with higher graduation rates tend to increase their graduation rates less rapidly over time. Nonetheless, it is important to note that the graduation rates did not take the enrollment size into account and were used to descriptively show the changes of graduation rates over a 6-year period. Therefore, the changes in unweighted graduation rates (i.e., small and large institutions are given equal weight) are not affected by differential changes in enrollment.

Table 2: Descriptive Statistics for All Variables

Variable	**Mean (SD)**						
Time-Invariant Variable							
Public	0.34 (0.47)						
Doctoral/Research	0.17 (0.38)						
Time-Varying Variable (by Year)							
Dependent Variable	*2011*	*2012*	*2013*	*2014*	*2015*	*2016*	*2017*
6-Year Graduation Rate	55.66	55.76	55.49	55.87	56.10	56.69	56.78
	(17.88)	(17.66)	(17.67)	(17.73)	(17.47)	(17.20)	(17.67)
Independent Variable	*2005*	*2006*	*2007*	*2008*	*2009*	*2010*	*2011*
FTE enrollment	5.66	5.72	5.82	5.94	6.10	6.22	6.29
(in 1000s)	(7.40)	(7.49)	(7.62)	(7.81)	(8.00)	(8.12)	(8.25)
% federal grant aid	31.00	29.12	29.53	30.75	30.83	36.88	40.36
	(15.87)	(15.94)	(16.42)	(16.43)	(16.24)	(17.21)	(17.16)
Instruction expenditures	9.17	9.74	9.44	9.52	9.85	10.04	9.93
	(7.39)	(15.15)	(7.67)	(7.90)	(8.44)	(8.20)	(8.00)
Research expenditures	1.83	1.82	1.78	1.81	1.89	2.04	2.05
	(6.67)	(6.83)	(6.69)	(6.85)	(7.34)	(7.62)	(7.78)
Academic support	2.31	2.46	2.41	2.50	2.55	2.61	2.56
expenditures	(2.83)	(3.96)	(2.91)	(3.45)	(3.37)	(3.25)	(3.19)
Student services	2.77	3.09	2.93	2.96	3.09	3.13	3.14
expenditures	(2.06)	(6.36)	(2.19)	(2.16)	(2.24)	(2.11)	(2.02)
Institutional support	4.19	4.87	4.40	4.45	4.55	4.53	4.42
expenditures	(3.55)	(18.08)	(3.99)	(3.88)	(3.85)	(3.80)	(3.58)

Note. N = 1,260; SD stands for standard deviation.

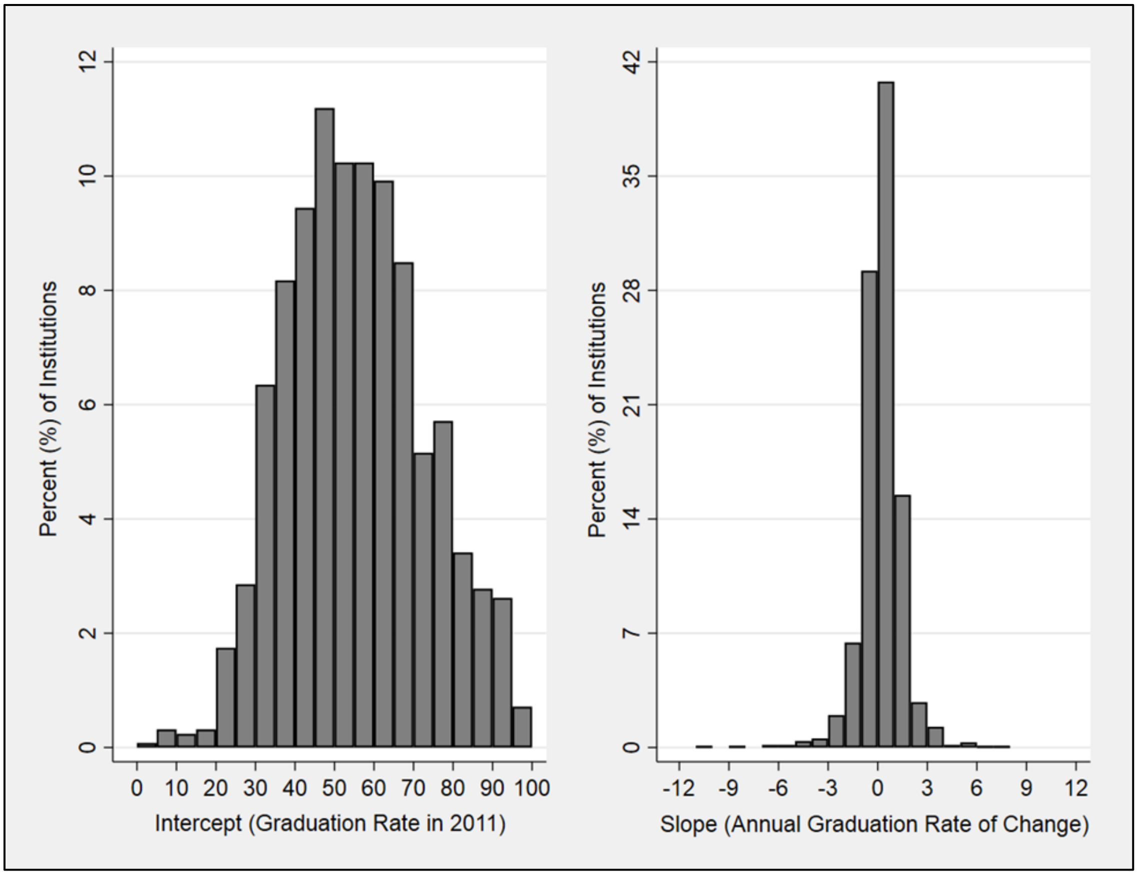

Figure 1: Distribution of Fitted Intercepts and Slopes from Separate OLS Regression Models

Table 3: Descriptive Statistics for Average of Fitted Separate Intercepts and Slopes

	Mean	**SD**	**Min**	**Max**
Initial status (intercept)	55.43	17.48	4.14	97.21
Rate of change (slope)	0.21	1.26	-10.11	7.79
Bivariate correlation between initial status and rate of change		-0.23		

Note. SD stands for standard deviation.

Unconditional Means Model & Unconditional Growth Model

To understand the total outcome variation, I tested the first model which is known as an unconditional means model with no predictors at either level. The unconditional means model is expressed by the following equations

Level 1: $Y_{ij} = ß_{0i} + \varepsilon_{ij}$

Level 2: $ß_{0i} = ß_{00} + U_{0i}$ (1)

where *i* denotes individual unit (i.e., institution) and *j* denotes time (i.e., year). In Level 1, $ß_{0i}$ represents initial level of graduation rates for individual institution *i* and the error term, ε_{ij} represents measurement error and unobserved heterogeneity. In Level 2, $ß_{00}$ represents average graduation rate at initial status (2011) and U_{0i} represents unexplained individual factors affecting graduation rate at initial status. Using an unstructured error covariance structure for the composite residual, I assumed each year point has its own variance in this unconditional means and following models.

As shown in Table 4, the fixed effect confirmed that the graduation rate grand mean across all years and institutions ($ß_{00}$) is non-zero. In addition, the results showed that the estimated within-institution variance (σ_{ε}^{2}) is 23.976; the estimated between-institution variance in initial status, is 286.168 (σ_{0}^{2}). Using the value of the two variances, I calculated the intraclass correlation coefficient (ICC) which describes the proportion of the total outcome variation that lies between institutions. Indicating that 92.3% of the total variation in graduation rate is attributable to differences among institutions, the intraclass correlation coefficient of 0.923 suggests that the between-institution variance account for the most of total variance in the sample. Finally, this unconditional means model is an improvement over a simple regression model because the log-likelihood (LR) test rejects the null hypothesis at p < .001 that random effect variance = 0. Therefore, I conclude that the unconditional means model provides a better fit than the linear model.

Table 4: Results of Fitting Unconditional Means and Unconditional Growth Models

	Parameter	Model 1	Model 2
Fixed Effects			
Initial Status			
Intercept	β_{00}	56.049*** (0.479)	55.425*** (0.486)
Rate of Change			
Intercept	β_{10}		0.208*** (0.036)
Variance Components			
Level 1			
Within-institution	σ_ε^2	23.976*** (0.390)	19.624*** (0.350)
Level 2			
In initial status	σ_0^2	286.168*** (11.538)	296.138*** (12.163)
In rate of change	σ_1^2		0.889*** (0.065)
Covariance	σ_{01}		-2.892*** (0.638)
Goodness-of-Fit			
Deviance		58624.58	58160.48
AIC		58648.58	58172.49
BIC		58669.84	58215.00

To evaluate the baseline amount of change, I then tested an unconditional growth model with time (i.e., year) as the only level-1 predictor and no substantive predictors at level 2. The unconditional growth model is expressed with the following equations:

Level 1: $Y_{ij} = \beta_{0i} + \beta_{1i} * Year_{ij} + \varepsilon_{ij}$

Level 2: $\beta_{0i} = \beta_{00} + U_{0i}$

$$\beta_{1i} = \beta_{10} + U_{1i} \quad (2)$$

where $ß_{1i}$ describes slope of change in six-year graduation rates for individual institution *i*. In Level 2, $ß_{10}$ represents graduation rate change by per unit of time (i.e., one year) and U_{0i} describes unexplained individual factors affecting change of graduation rates.

Model 2 in Table 4 presents the results of fitting the unconditional growth model to the graduation rate data. The fixed effects, $ß_{00}$ and $ß_{10}$, show that the average true change trajectory for graduation rate has a non-zero intercept of 55.425 and a non-zero slope of 0.208. To assess whether there is hope for future analyses – whether there is statistically significant variation in individual initial status or rate of change that level-2 predictors could explain – I examined the variance components. Comparing the level-1 variance within institutions (σ_{ε}^{2}) in Model 2 to that of Model 1, I found a decline of 0.182 (from 23.976 to 19.624). In other words, 18.2% of the within-institution variation in graduation rates in is systematically associated with linear time (i.e., year).

The level-2 variance components quantify the amount of unpredicted variation in the individual growth parameters. σ_0^2 and σ_1^2 represent the unconditional variation in initial status and in rates of change, respectively. Suggesting that it may be worthwhile using level-2 predictors to explain heterogeneity in the random intercept and slope, the results show that there is non-zero variability in both initial status and rate of change. The covariance (σ_{01}) of the level-2 residuals indicates that institutions with higher graduation rate increase their graduation rates less rapidly over time because the covariance showed a negative relationship between initial status and change of graduation rates. Finally, this unconditional growth model offers an improvement over the unconditional means model because the LR test (i.e., the difference in deviance statistics) allows me to reject the null hypothesis at the $p < .001$ level that all coefficients of three new added parameters are simultaneously 0. Therefore, I conclude that the unconditional growth model provides a better fit than the unconditional means model.

Multilevel Models for Change with Covariates

Model 3 includes *Public* as a time-invariant predictor which recodes an individual's static status (Singer & Willett, 2003). The conditional growth model with the time-invariant predictor of Public is expressed with the following equations:

Level 1: $Y_{ij} = ß_{0i} + ß_{1i} * Year_{ij} + \varepsilon_{ij}$

Level 2: $ß_{0i} = ß_{00} + ß_{01} * Public_{i} + U_{0i}$

$$ß_{1i} = ß_{10} + ß_{11} * Public_{i} + U_{1i} \tag{3}$$

where $ß_{00}$ represents average of graduation rates in 2011 for private, not-for-profit institutions and $ß_{10}$ represents average annual rate of change in graduation rates for private, not-for-profit institutions. $ß_{01}$ describes average difference in graduation rates in 2011 between public and private, not-for-profit institutions and $ß_{11}$ describes average difference in change of graduation rates at between public and private, not-for-profit institutions by per unit of time (i.e., one year) U_{0i} describes unexplained individual factors affecting change of graduation rates.

The results indicate that the estimated difference in initial graduation rates (in 2011) between public and private, not-for-profit institutions is -9.179. In other words, the initial graduation rates in 2011 are about 9.2% lower for public institutions than for private, not-for-profit institutions. In contrast, the estimated differential in the rate of change in graduation rates between public and private, not-for-profit institutions shows public institutions increased graduation rates over time more rapidly than private, not-for-profit institutions. Adding the predictor of *Public* significantly influenced the rate of change. The level-2 intercept in rate of change is no longer significant and suggested that the slope of the population average change trajectory in graduation rates is zero after controlling for institutional type (i.e., Public versus Private, not-for-profit). The level-2 predictor partially impacted variability. The within-institution variance component for Model 3 is identical to that of Model 2 and suggests the need to explore the effects of time-varying predictors. In contrast, the level-2 predictor impacted all level-2 variance components and reduced the variability. σ_0^2 and σ_1^2 declined by 7.5% and 4.1%, respectively from Model 2. Because all level-2 variances are still statistically significant, potentially explainable residual variations in initial status, change and covariance remains. In particular, the results suggest that the continued presence of potentially explainable residual variation in rates of change.

Adding another time-invariant predictor of both initial status and change, Model 4 evaluates the effects of *Public* and *Doctoral/Research* on initial status and rate of change in graduation rates after controlling for the effects of each other on initial status and rate of change. The conditional growth model with two time-invariant predictors is expressed with the following equations:

$$\text{Level 1:} \quad Y_{ij} = ß_{0i} + ß_{1i} * Year_{ij} + \varepsilon_{ij}$$

$$\text{Level 2:} \quad ß_{0i} = ß_{00} + ß_{01} * Public_i + ß_{02} * Doctoral/Research_i + U_{0i}$$

$$ß_{1i} = ß_{10} + ß_{11} * Public_i + ß_{12} * Doctoral/Research_i + U_{1i} \qquad (4)$$

where $ß_{01}$, $ß_{02}$, $ß_{11}$, and $ß_{12}$ describe controlled effects of *Public* and *Doctoral/Research*.

This study found that the initial graduation rates in 2011 are about 17% higher for doctoral/research institutions which have more research-oriented missions than for non-doctoral/research institutions. In addition, the estimated differential in the rate of change in graduation rates between them shows doctoral/research institutions increased graduation rates over time more rapidly. The magnitude of the fixed effect of *Public* on initial status increased by 43.3% from -9.179 to -13.243, suggesting that the differential between public and private, not-for-profit institutions in initial status of graduation rates in 2011 became larger after controlling for Carnegie classification. Regarding variance components, the results suggest that the additional level-2 predictor of Carnegie classification to Model 3 considerably decreased variability in initial status by 13.7% from 277.174 to 239.192

whereas variability in rate of change remained stable. The level-1 variance component remained stable because time-invariant predictors cannot explain much within-institution variation (Singer & Willett, 2003). Thus, the statistically significant within-institution variance component (σ_ε^2) indicates a need to explore the effects of time-varying predictors. Finally, I tested whether this growth model with two time-invariant predictors offers an improvement over the unconditional growth model. Because the LR test (i.e., the difference in deviance statistics) allowed me to reject the null hypothesis at the $p < .001$ level that all coefficients of the new added parameters are simultaneously 0. Therefore, I conclude that the growth model with two predictors provides a better fit than the unconditional growth model.

To investigate the effects of time-varying predictors, Model 5 includes five forms of institutional expenditures, enrollment size, and percentage of students who received federal aid in addition to two time-invariant variables such as *Public* and *Doctoral/Research*. Unlike the time-invariant variables, the time-varying predictors whose values differ over time record an individual's potentially differing status on each associated measurement occasion (Singer & Willett, 2003). In this study, I restricted attention to examining the main effects of time-varying predictors in order to avoid the complexity of analysis and interpretation in the model, assuming that effects of time-varying predictors do not vary over time. My extended analysis excluded in this study supported the assumption, showing that the effects of all time-varying predictors except institutional support expenditures do not vary over time. The full growth model is expressed with the following equations:

Level 1: $Y_{ij} = ß_{0i} + ß_{1i}*Year_{ij} + + ß_{20\text{-}26}*Z_{ij}$ *(FTE enrollment, % federal grant aid, five forms of institutional expenditures)* $+ \varepsilon_{ij}$

Level 2: $ß_{0i} = ß_{00} + ß_{01}*Public + ß_{02}*Doctoral/Research + U_{0i}$

$$ß_{1i} = ß_{10} + ß_{11}{*Public} + ß_{12}{*Doctoral/Research} + U_{1i} \quad (5)$$

where $ß_{20}$–$ß_{26}$ describe the estimated effects of seven time-varying predictors.

Regarding the effects of time-varying predictors, the results in Model 5 suggest that the institutions with larger enrollments had higher graduation rates while the percentage of students receiving federal grant aid was negatively associated with six-year graduation rates. In addition, the results suggest that institutional expenditures on instruction, research, and academic support increased graduation rates, whereas institutional support expenditures decreased graduation rates and student services expenditures did not influence graduation rates. The results provide empirical evidence on how the time-varying variables impact graduation rates. Nonetheless, it is noteworthy to note that recentering time-varying predictors such as centering on mean, or other meaningful constant, can offer more meaningful interpretation at a certain value that we may be interested in while centering does not impact the fit of the model and conclusion on findings (Singer & Willett, 2003).

The inclusion of time-varying predictors also changed the magnitude of variance components in Model 5. Unlike time-invariant predictors which do not impact within-institution variance, time-varying predictors can influence all three variance components (i.e., within-institution variance, variance in initial status, and variance in rate of change) because the predictors can vary within and between institutions (Singer & Willett, 2003). Adding time-varying predictors to the growth model with two time-invariant predictors (Model 4) reduced the magnitude of the initial status variance component by 19.1% from 239.192 to 193.506 whereas the other variance components remained relatively stable. Finally, Model 5 offers an improvement over the growth model with two time-invariant predictors (Model 4) because the LR test (i.e., the difference in deviance statistics) allows me to reject the null hypothesis at the $p < .001$ level that all three new added parameters are simultaneously 0. Therefore, I conclude that the full growth model provides a better fit than the growth model with two time-invariant predictors.

Table 5: Results of Fitting Conditional Growth Models with Covariates

	Parameter	Model 3	Model 4	Model 5 (Full)
Fixed Effects				
Initial Status				
Intercept	β_{00}	58.565*** (0.588)	56.985*** (0.559)	55.961*** (0.612)
Public	β_{01}	-9.179*** (1.005)	-13.243*** (0.980)	-15.560*** (0.975)
Doctoral/Research	β_{02}		17.010*** (1.225)	6.266*** (1.420)
Rate of Change				
Intercept	β_{10}	0.066 (0.043)	0.047 (0.044)	0.117* (0.046)
Public	β_{11}	0.415*** (0.074)	0.367*** (0.077)	0.247** (0.078)
Doctoral/Research	β_{12}		0.202* (0.097)	-0.058 (0.099)
FTE enrollment	β_{20}			0.606*** (0.068)
% federal grant aid	β_{21}			-0.055*** (0.007)
Instruction expenditures	β_{22}			0.193*** (0.043)
Research expenditures	β_{23}			0.126* (0.052)
Academic support expenditures	β_{24}			0.343*** (0.070)
Student services expenditures	β_{25}			-0.107 (0.094)
Institutional support expenditures	β_{26}			-0.156*** (0.037)

Table 5 (continued)

	Parameter	Model 3	Model 4	Model 5 (Full)
Variance Components				
Level 1				
Within-institution	σ_ε^2	19.624***	19.623***	19.895***
		(0.350)	(0.350)	(0.359)
Level 2				
In initial status	σ_0^2	277.174***	239.192***	193.506***
		(11.407)	(9.894)	(8.521)
In rate of change	σ_1^2	0.851***	0.845***	0.852***
		(0.063)	(0.063)	(0.064)
Covariance	σ_{01}	-2.035***	-2.487***	-3.111***
		(0.606)	(0.568)	(0.527)
Goodness-of-Fit				
Deviance		58065.54	57860.26	57661.94
AIC		58081.55	57880.26	57695.94
BIC		58138.23	57951.11	57816.38

Alternative Error Covariance Structure for Heteroscedasticity and Autocorrelation

Two features of multilevel models are autocorrelation and heteroscedasticity within level-2 units of analysis (Singer & Willett, 2003). Figure 2 shows the empirical autocorrelation with the graduation data, indicating temporal dependency of composite residuals. The previous models in this study rejected the OLS assumptions such as independence of errors and accepted an unstructured error covariance structure for the composite residual, assuming each time point has its own variance (i.e., 10 unknown parameters – four variances and 6 covariances) in a model. The unstructured error covariance structure strongly allowed for heteroscedasticity and autocorrelation among the composite residuals, but a more parsimonious structure might be more appropriate to reduce “wasting” considerable degrees of freedom (Singer & Willett, 2003). Therefore, I sought to explore what type of correlation structure can be suggested to allow for heteroscedasticity and autocorrelation.

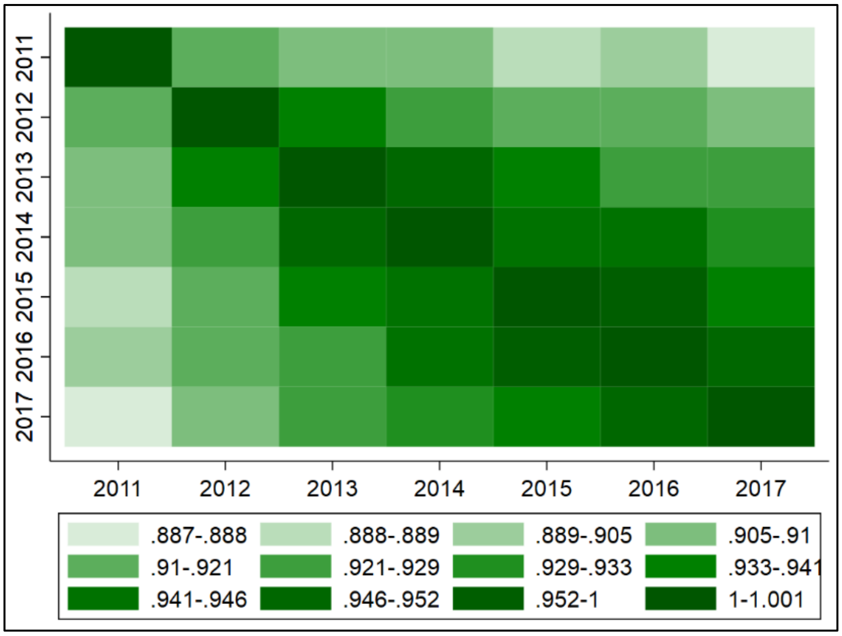

Figure 2: Empirical Correlations of Composite Residuals between Years

Table 6 presents results on the value of variance components and goodness-of-fit. In order to determine whether there is more appropriate alternative error covariance structure, Model 6 and Model 7 particularly used first-order autoregressive and Toeplitz covariance structures for model errors, respectively. By using AIC and BIC to gauge among competing models, which are not nested within each other, the results suggest that although the alternative structure appeared to provide very small improvement on the multilevel model for change with an unstructured error covariance structure, first-order autoregressive structure may be the most appropriate model among three alternatives including the unstructured error covariance structure. However, I acknowledge that it is entirely possible that there are other error structures that would be superior for this data. Figure 3 presents the fitted residual autocorrelations under the first-order autoregressive error covariance matrix and clearly shows that Model 6 fits well. Therefore, I chose Model

6 with the first-order autoregressive error covariance structure because Model 6 has the most parsimonious structure but better fit than other alternatives presented in this study.

Table 6: Variance and Goodness-of-Fit by Model with Alternative Error Covariance Structure

	Parameter	Model 6	Model 7
Variance Component			
In initial status	σ_0^2	182.24***	184.005***
In rate of change	σ_1^2	0.521***	0.610***
Residual variance	σ_ε^2	22.827***	21.736***
Autocorrelation between years	ρ	0.185***	
Covariance	σ_{01}		2.924***
Goodness-of-Fit			
Deviance		57650.76	57666.48
AIC		57684.75	57700.48
BIC		57805.20	57820.93

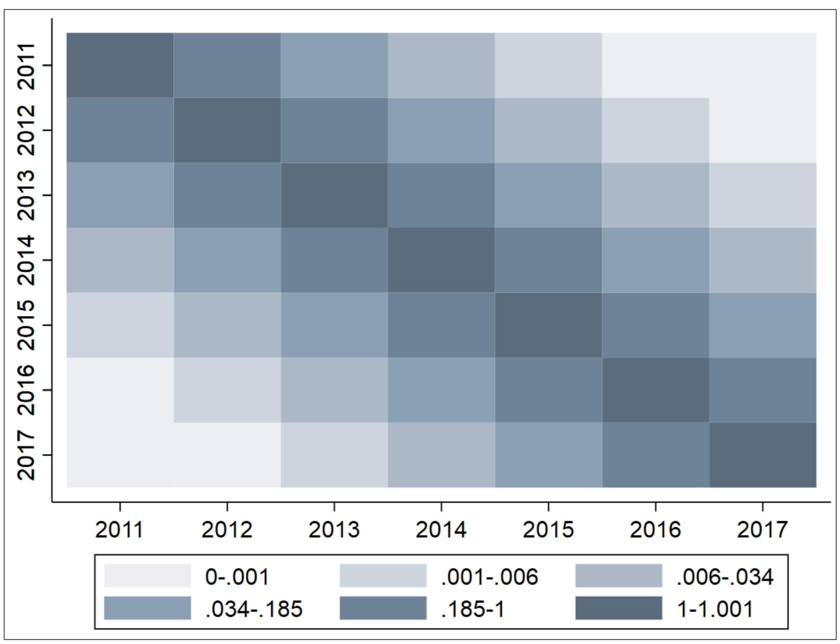

Figure 3: Fitted Correlations of Composite Residuals between Years under First-Order Autoregressive Covariance Structure

Building Final Model

Using literature review, the four-step process using Singer and Willett's (2003) models, and error covariance structure to address heteroscedasticity and autocorrelation in the longitudinal data, I build the final model which fits best to the data as presented in Table 7. In addition, the final model excludes predictors which have no effects on initial status and rate of change regarding graduation rates to build a parsimonious model that accomplishes a desired level of explanation with as few predictor variables as possible.

Table 7: Results of Fitting the Final Growth Model to the IPEDS Data

	Parameter	Final Model	
		Coefficient	**SE**
Fixed Effects			
Initial Status			
Intercept	β_{00}	55.806***	0.600
Public	β_{01}	-15.459***	0.959
Doctoral/Research	β_{02}	6.314***	1.420
Rate of Change			
Intercept	β_{10}	0.100*	0.044
Public	β_{11}	0.233**	0.074
FTE enrollment	β_{20}	0.606***	0.069
% federal grant aid	β_{21}	-0.052***	0.007
Instruction expenditures	β_{22}	0.175***	0.042
Research expenditures	β_{23}	0.133*	0.053
Academic support expenditures	β_{24}	0.333***	0.072
Institutional support expenditures	β_{26}	-0.176***	0.032
Variance Components			
Level 1			
Within-institution	σ_{ε}^2	22.823***	0.612
Autocorrelation between years	ρ	0.185***	0.020
Level 2			
In initial status	σ_0^2	182.104***	8.142
In rate of change	σ_1^2	0.520***	0.070
Goodness-of-Fit			
Deviance		57646.36	
AIC/BIC		57676.35/57782.62	

Note: SE stands for standard error.

Chapter Four: Conclusion

In this study, I utilized data on institutional graduation rates from IPEDS during the period of 2011 to 2017 to answer to the research question: what institutional characteristics are associated with the patterns of change in graduation rates? The results suggest that public institutions had lower graduation rates in 2011, but they increased more sharply than private, not-for-profit institutions during the period of observations. Although doctoral/research institutions with research-oriented missions had higher graduation rates in 2011, no difference was found in the change of graduation rates between doctoral/research institutions and other institutions. While enrollment size had a positive relationship with institutional graduation rates, the percentage of federal grant recipients was negatively associated with graduation rates. All expenditures except institutional support were positively associated with institutional graduation rates. In contrast, institutional support expenditures are negatively associated with institutional graduation rates. Although this study illuminates a significant relationship between institutional characteristics and graduation rates, future research could focus on causal effects of the variables on graduation rates.

This study also sought to build a statistical model which fits the longitudinal data based on prior research, model testing, and comparison of model fit. Testing goodness-of-fit through multilevel modeling and first-order autoregressive error covariance structure provided improvement over the traditional approaches which do not allow for specific temporal patterns of autocorrelation and heteroscedasticity. However, there are two limitations to be considered in further analysis. First, due to missing data due to deletions, the data used in this analysis was unbalanced. I acknowledge that the results might be biased due to losing data. Therefore, it may be necessary to apply statistical approaches to deal with missing data to incorporate more information about the population. In addition, the results suggested that the time-invariant predictor indicating public versus private, not-for-profit institutions significantly impact the level-2 variance and reduced the impact of time on change in graduation rates. Therefore, it may be necessary to examine heterogeneous variance components that differentiate between public and private, not-for-profit institutions. Lastly, as supported in previous studies, difference in enrollment size of institutions matters. Although enrollment size is added into the model, it may be better to apply weighting to adjust the difference in size of institution.

References

Burke, J. C., & Minassians, H. (2002). Reporting indicators: What do they indicate? In J. Burke & H. Minassians (Eds.), Reporting higher education results: Missing links in the performance chain. *New Directions for Institutional Research, 116*, 33–58.

Christal, M. E. (1998). State survey on performance measures: 1996–97. Denver, CO: State Higher Education Executive Officers/Education Commission of the States.

Cragg, K. M. (2009). Influencing the probability for graduation at four-year institutions: A multi-model analysis. *Research in Higher Education, 50*(4), 394–413.

Gansemer-Topf, A. M., & Schuh, J. H. (2003). Instruction and academic support expenditures: An investment in retention and graduation. *Journal of College Student Retention: Research, Theory & Practice, 5*(2), 135–145.

Gansemer-Topf, A. M., & Schuh, J. H. (2006). Institutional selectivity and institutional expenditures: Examining organizational factors that contribute to retention and graduation. *Research in Higher Education, 47*(6), 613–642.

Gross, J. P., & Berry, M. S. (2016). The relationship between state policy levers and student mobility. *Research in Higher Education, 57*(1), 1–27.

Gross, J. P., Torres, V., & Zerquera, D. (2013). Financial aid and attainment among students in a state with changing demographics. *Research in Higher Education, 54*(4), 383–406.

Goenner, C. F., & Snaith, S. M. (2004). Predicting graduation rates: An analysis of student and institutional factors at doctoral universities. *Journal of College Student Retention: Research, Theory & Practice, 5*(4), 409–420.

Ishitani, T. T. (2006). Studying attrition and degree completion behavior among first-generation college students in the United States. *Journal of Higher Education, 77*(5), 861–885.

National Research Council. (2013). Improving measurement of productivity in higher education. National Academies Press.

Oseguera, L. (2005). Four and six-year baccalaureate degree completion by institutional characteristics and racial/ethnic groups. *Journal of College Student Retention: Research, Theory & Practice, 7*(1), 19–59.

Sanford, T., & Hunter, J. M. (2011). Impact of Performance Funding on Retention and Graduation Rates. *Education Policy Analysis Archives, 19*, 33.

Schneider, M., & Yin, L. (2011). The hidden costs of community colleges. American Institutes for Research.

Shin, J. C. (2010). Impacts of performance-based accountability on institutional performance in the US. *Higher Education, 60*(1), 47-68.

Singer, J. D., Willett, J. B., & Willett, J. B. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. Oxford University Press.

Tandberg, D. A., & Hillman, N. W. (2014). State higher education performance funding: Data, outcomes, and policy implications. *Journal of Education Finance, 39*(3), 222–243.

Titus, M. A. (2006a). No college student left behind: The influence of financial aspects of a state's higher education policy on college completion. *Review of Higher Education, 29*(3), 293–317.